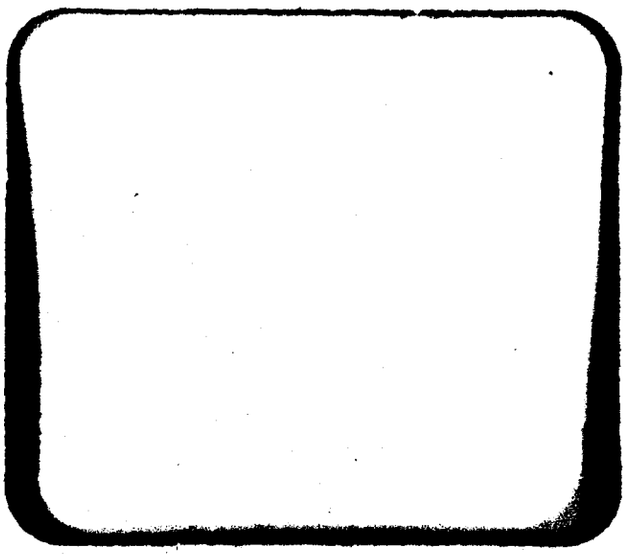
ORGANIZATIONAL RESEARCH

DEPARTMENT OF PSYCHOLOGY • UNIVERSITY OF WASHINGTON, SEATTLE, WASHINGTON



ORGANIZATIONAL RESEARCH

DEPARTMENT OF PSYCHOLOGY

UNIVERSITY OF WASHINGTON

SEATTLE, WASHINGTON

(1)

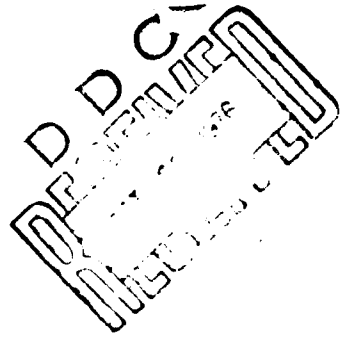
INFORMATION RELEVANCE, CONTENT, AND SOURCE CREDIBILITY IN THE REVISION OF OPINIONS[1]

Lee Roy Beach[2], Terence R. Mitchell,
Marcia D. Deaton, and Joyce Prothero

University of Washington
Seattle, Washington

Technical Report 76-80

March 1976



Project Number BSA 936
Battelle Institute Northwest

Contract N00014-76-C-0193
Office of Naval Research

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: TR-76-8Ø	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): INFORMATION RELEVANCE, CONTENT AND SOURCE CREDIBILITY IN THE REVISION OF OPINIONS	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
6. PERFORMING ORG. REPORT NUMBER: 76-2814	
7. AUTHOR(s): Lee Roy Beach, Terence R. Mitchell, Marcia D. Deaton, Joyce Prothero	8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0193
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Organizational Research, Department of Psychology (NI-25), University of Washington, Seattle, WA 98195	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: BSA-936
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research Code 452, 800 North Quincy Street, Arlington, VA 22217	12. REPORT DATE: March 1976
	13. NUMBER OF PAGES: 26
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 28 p.	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Relevance, Content, Source Credibility, Reliable, discounted, Revise Opinions, Probability of Success, Acceptability

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In the first of two experiments participants assessed either (1) the probability of success or (2) the acceptability of hypothetical job candidates. These judgments were made on the basis of information about particular traits or abilities possessed by the candidate. The traits themselves varied in relevance to the job and the sources of information differed in credibility. Results showed that the information was 'discounted' (had reduced influence

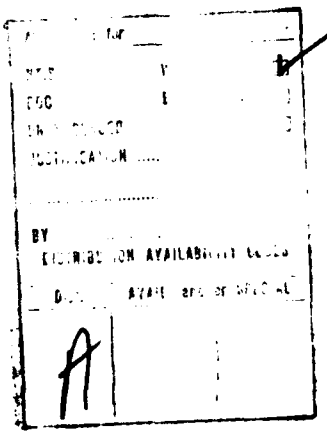
DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

387783

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. on the judgment) when its relevance decreased and when its sources' credibility decreased. Four discounting models were examined and rejected, although by estimating parameters from part of the data and fitting the remainder, one model could be acceptably modified to account for the data. Discounting was a multiplicative rather than a subtractive process and the sameprocess apparently accounted for both the acceptability judgments and the probability assessments, which were virtually identical. In the second experiment, participants received three pieces of information about the potential negative effects of constructing nuclear power plants on a number of potential sites. For each site, participants revised their judgments about the favorableness of the site after each of the three pieces of negative information. Again, discounting was obtained as a function of the credibility of the information sources and a form of 'conservatism' was obtained; regardless of the source and the content of the message, the amount of revision decreased from the first to the second to the third revision.



Information Relevance, Content, and Source Credibility In the Revision of Opinions[1]

There has been of late a good deal of interest in how people utilize information to revise their opinions about the truth of hypotheses when the information source, and therefore the information itself, is not wholly credible or reliable. The focus has been upon the degree to which people's revisions of their subjective probabilities conform to the prescriptions of the Bayesian Hierarchical Inference model, the normative model for revising prior opinion in light of the implications of new data when the implications have been discounted to allow for the lack of credibility of the source (Gettys & Wilke, 1969; Peterson, 1973; Schum & DuCharme, 1971).

Formally, the Bayesian model begins with prior opinion being represented as a ratio of the probability that the hypothesis in question is true to the probability that it is not, $\frac{P(H)}{P(\overline{H})}$. After observation of a datum (information) this prior probability ratio, Ω_0, is multiplied by a discounted likelihood ratio, Λ, which incorporates the probability that the datum is accurate and what it means for the hypothesis if it is and the probability that the datum is not accurate and the implication of that for the hypothesis,

$$\Lambda = \frac{P(D^*|H)}{P(D^*|\overline{H})} = \frac{P(D^*|D)\ P(D|H) + P(D^*|\overline{D})\ P(\overline{D}|H)}{P(D^*|D)\ P(D|\overline{H}) + P(D^*|\overline{D})\ P(\overline{D}|\overline{H})}$$

where D^* is the reported datum, D and $\overline{D}$ are the truth or falsity of the report, and H and $\overline{H}$ are the truth or falsity of the hypothesis. The revised opinion is a ratio of the posterior probabilities, $\Omega_1 = \frac{P(H|D^*)}{P(\overline{H}|D^*)} = \Lambda\ \Omega_0$ (Schum & DuCharme, 1971).

Research comparing unaided human opinion revision with the prescriptions of the model suggests that people do indeed discount noncredible data but the normative model probably is not an accurate description of how they do it (e.g., Peterson, 1973). An alternative, descriptive, model is in order. Two simple, straightforward descriptive alternatives have been suggested, one by Snapper and Fryback (1971) and another by Gettys, Kelly, and Peterson (1973). The Snapper and Fryback model hypothesizes that people use a discounted likelihood ratio in their revisions, but that instead of being the complicated discounting procedure prescribed by the Bayesian model, it is merely the likelihood ratio, L, that they would use if the source were 100% credible multiplied by the credibility, r, of the source, i.e., $\Omega_1 = (Lr)\ \Omega_0$. The Gettys et al. model states that people revise as though the source were 100% credible and then they reduce the posterior probability ratio by some amount in order to reflect their lack of confidence in the data--the amount of reduction is determined by multiplying the posterior by (perhaps, but not necessarily) the source's credibility, r, i.e., $\Omega_1 r = L\Omega_0$. Supporting evidence for the two models is presented in the articles in which they were proposed.

The major difference between these two models is the hypothesized order in which revising and discounting operations take place and the locus of the discounting. Snapper and Fryback hypothesize that the likelihood ratio is discounted first and that it then is used in the revision. Gettys et al. hypothesize that the revision takes place first and then the posterior opinion is discounted. Additionally, Snapper and Fryback specify that the "discounting factor", as we shall call it, is the source's credibility, r, while Gettys et al. do not make this restriction--although their data suggest it. For ease

of exposition we will henceforth refer to these two models as the "r-multiplicative models".

There is another plausible simple descriptive model for discounting--one that is a variation of the previous two. It is that instead of the discounting process being multiplicative, it is subtractive. That is, people reduce either the likelihood ratio or the posterior probability ratio by merely subtracting some amount that reflects the degree of non-credibility of the information. We will call this the subtractive model. The first purpose of the present research is to examine the adequacy of these discounting models: Is discounting multiplicative or subtractive and if it is the former, is r the multiplicative factor?

We should also point out that there is a large body of research in social psychology which is conceptually related to the question of how people revise their opinions, the research on attitude change. In this work the focus has been on the conditions that promote more or less attitude change rather than upon the congruence of the change with the prescriptions of a normative model. However, if attitudes are opinions just as subjective probabilities are, it seems reasonable to speculate that similar mechanisms might underlie the revisions of both. To see if this is indeed the case, attention should be directed to exploring the similarities among different kinds of opinion revisions with the goal of obtaining a broader view of the revision process. The second purpose of the present research is to make a beginning in this direction.

Experiment 1

In this first experiment the participants were asked to evaluate hypothetical job candidates. There were two experimental conditions, each with a separate group of participants. In the first, participants assessed the "probability of success" on a particular job for each candidate, and in the second condition, participants rated the "acceptability" of each candidate for the particular job. This second condition yields evaluative responses of the kind traditionally obtained in attitude research.

Both conditions used identical sets of information. Each candidate for a particular job was described as having some personality trait or characteristic that had been assessed by some outside agency. This trait varied in its relevance to the job and the agencies varied in their credibility, where the latter was defined as the percentage of times in the past that the source had provided accurate information. Relevance and credibility are variables that have been shown to be important in attitude change research (e.g., Hovland and Weiss, 1951; Hill, 1963; Choo, 1964).

Method

Experimental materials consisted of hypothetical "referral reports" in which applicants for one of three jobs (Driver Education Teacher, Journalist, or Social Worker) were described as having an attribute that was of one of three degrees of relevance (high, medium, or low) to job performance. The referral reports supposedly came from personnel agencies (sources) that had differing degrees of credibility as a result of their past records of accurate evaluations (100%, 75%, 50%, 25%). The 3 jobs x 3 relevance levels x 4

credibility levels = 36 hypothetical referral reports, 1 for each of 36 hypothetical applicants. Each report constituted a page (Fig. 1) in a booklet that was given to each participant in the experiment; the pages of each booklet were scrambled to control for presentation order effects.

Insert Figure 1 about here

The three levels of relevance of the applicant attributes for each job were, for the Driver Education Teacher, (high) Perceptual accuracy--speed and depth perception, (medium) Ability to solve abstract theoretical problems, and (low) Artistic ability; for the Journalist, (high) Verbal ability, (medium) Arithmetic ability, and (low) Musical ability; for the Social Worker, (high) Sensitivity and compassion, (medium) Memory--especially for numbers, and (low) Knowledge of art history. The hypothetical applicants all were given favorable evaluations in the reports in order to hold favorableness constant.

Selection of the attributes and specification of the relevance levels were determined before the experiment. Twnety persons (from the same population as the experimental participants) used 5-point scales to rate a large number of personal attributes in terms of their importance for performance of each of the three jobs. The three attributes that received the greatest level of agreement about being of high, medium and low importance (relevance) for a job were selected for use in the experiment. Then, in the experiment proper, participants made similar ratings at the end of the session. In cases where individual participants gave reversed judgments (e.g., they rated a characteristic as low that the earlier group had rated as high), their

Candidate # 1601

Job applied for Journalist

Evaluation very high in verbal ability

Source AAA Employment Agency

Source reliability correct about one-fourth of the time.

Acceptability of candidate

Figure 1

responses were analyzed in terms of their own judgments.

Probability Condition. In the first condition of the experiment, the participants read each referral report and estimated the probability that the applicant would succeed in the job. Estimates were made by marking a 100-point scale that was printed on the page below each report. The instructions explained that .0 on the scale meant that the applicant definitely would fail, 1.00 meant that he or she definitely would succeed, .50 meant that one really could not tell one way or the other, and that in-between numbers meant in-between degrees of certainty of success or failure. Thirty-four college students were paid $2.00 each for participation.

Acceptability Condition. The procedure for the second condition of the experiment was the same as for the first except that the participants were asked to use the 100-point scales to evaluate the acceptability of each applicant for the job. The 0 point was to indicate total unacceptability and 100 indicated total acceptability. Thirty-six college students were each paid $2.00 for participating.

Predictions. On the basis of the attitude change literature, it was expected that the less relevant the information was to the occupations in question, the more it would be discounted and that the assessed probability of success and the judged acceptability of the applicant would decrease as a function of decreased relevance. And, on the basis of both the attitude change literature and the decision literature it was expected that, for each level of relevance of the information, assessed probability and judged acceptability would decrease as a function of decreased source credibility. Moreover, if there is a mechanism for discounting irrelevant and/or unreliable information,

and if it is common to different kinds of dependent variables, then the results for assessed probabilities and judged acceptability would be expected to be highly similar.

The first test of the descriptive models consists of seeing whether discounting is subtractive or multiplicative using Anderson's analytic techniques (1971). These involve both graphing the data for visual inspection and subjecting them to analyses of variance. If discounting is subtractive the visual inspection should show parallel lines. The main effects of the analysis of variance should be significant but the relevant interaction should be nonsignificant. If discounting is multiplicative the visual inspection should show converging or diverging lines. The main effects of the analysis of variance should be significant and the relevant interaction also should be significant. If the process is purely multiplicative with no other component, reanalysis of the data under a log transform should eliminate the previously significant interaction.

If discounting proves to be multiplicative, the next step is to see if source credibility, r, is the proper multiplicative discounting factor. A test of this is afforded by the fact that for some of the hypothetical applicants the sources were represented as 100% credible. Participants' probability assessments for these cases can be regarded as a revision of some unknown prior opinion using some unknown but undiscounted likelihood to obtain the known (the observed assessments) undiscounted posterior opinions. That is, in the equations for the r-multiplicative models r = 1.00. As can be seen from the equations presented earlier for both of these models, when r = 75%, 50%, or 25% the observed assessments should be 75%, 50%, or 25% of what they

are for 100%. So, for a given level of relevance, knowing the mean observed assessment for when the source is 100% reliable makes it possible to compute what the assessment ought to be for the other three levels of reliability if either of the two multiplicative models is correct.

Results

Probability Condition. In Fig. 2, for each job the mean probability assessments are plotted against information relevance for each level of source credibility and in Fig. 3 they are replotted with source credibility as the independent variable. Looking at Fig. 3, if the influence of the information on assessed probability decreases with decreasing relevance the curves for the highly relevant information should be highest, the curves for the low relevance information should be lowest, and the medium relevance curve should be between the other two; which is indeed the obtained order although there are small differences in the heights of the curves for the different jobs. If information is discounted as a function of source credibility the curves should decline from left to right, which they do.

Insert Figures 2 & 3 about here

If discounting due to source credibility is subtractive, the curves should be parallel and if it is multiplicative, the curves should converge from left to right; they appear to converge but it is difficult to tell merely by visual inspection.

The assessments were subjected to a 3 x 3 x 4 (jobs x relevance x source credibility) analysis of variance. There was no significant main effect for

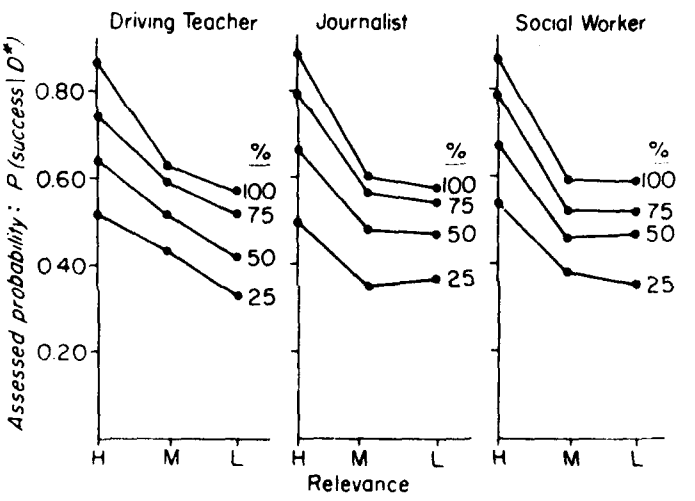

Driving Teacher
Journalist
Social Worker
Assessed probability: P(success | D*)
0.80
0.60
0.40
0.20
%
100
75
50
25
H
M
L
Relevance

Figure 2

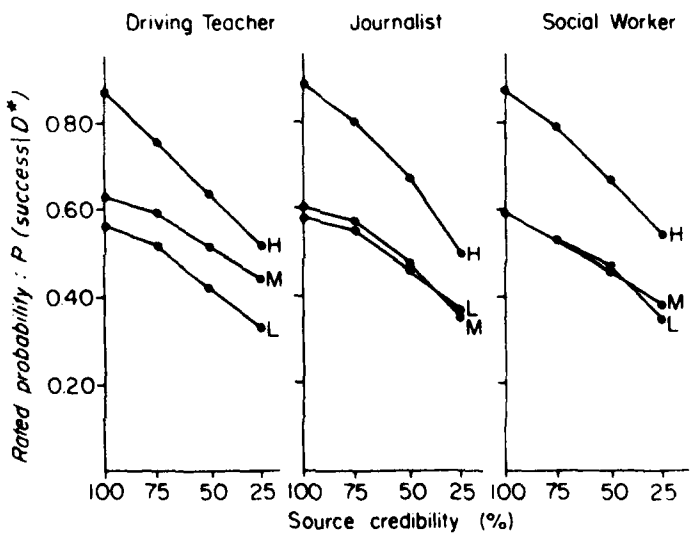

Driving Teacher
Journalist
Social Worker
Rated probability : P (success | D*)
0.80
0.60
0.40
0.20
H
M
L
100 75 50 25
Source credibility (%)

Figure 3

jobs, $F(2,66) = .148$, but then none was expected. There was a significant jobs x relevance interaction, $F(4,132) = 4.83$, $p < .001$, resulting from the experimenters' failure to make the three levels of relevance precisely the same for all three jobs (see Fig. 3); the interaction is of no substantive interest.

The important results are the significant main effects of relevance, $F(2,66) = 87.31$, $p < .001$, and source credibility, $F(3,99) = 82.43$, $p < .001$, and the relevance x source credibility interaction, $F(6,198) = 8.43$, $p < .001$.

The main effect for relevance confirms that the left-to-right decreases of the curves in Fig. 2 and the differences in the elevations of the curves in Fig. 3 are statistically significant, which means that information is discounted more as its relevance decreases. The main effect for source credibility confirms that the differences in the elevations of the curves in Fig. 2 and the left-to-right decreases of the curves in Fig. 3 are statistically significant, which means that information is discounted more as its source's credibility decreases. The significant interaction confirms that the curves in both figures are not parallel, which means that discounting is multiplicative, not subtractive.

To assure that the relevance x source credibility interaction is solely the result of multiplicative discounting the data were transformed into logs and the analysis of variance was done again. In logs multiplication become addition and the interaction should no longer be significant. As with the untransformed data, there was no significant main effect for jobs, $F(2,66) = .11$, and there was a significant jobs x relevance interaction, $F(4,132) = 3.68$, $p < .025$. Relevance remained a significant main effect, $F(2,66) =$

61.59, $p < .001$, as did source credibility, $F(3,99) = 64.54$, $p < .001$. But, the relevance x source credibility interaction was no longer significant, $F(6,198) = 1.07$. Thus, we can feel fairly confident that this interaction was attributable to a multiplicative discounting process.

Having eliminated the subtractive model, we turn to the evaluation of the r-multiplicative models. To do this the participants' probability assessments were converted to ratios and averaged across jobs for each of the three levels of relevance and each of the four levels of source credibility. These means are the observed data that are connected by the solid lines in panel a of Fig. 4. Then for each level of relevance, the mean assessment for 100% credible sources was multiplied by .75, .50, and .25; these are the predicted points connected by the dashed lines in panel a of Fig. 4. If either of the r-multiplicative models is correct in asserting that discounting consists merely of multiplying the likelihood ratio or the ratio of posterior probabilities by the credibility of the source, the two sets of points should correspond for each level of relevance, which they do not. If the models are correct the obtained curves should all be straight but the curve for the high relevance information clearly is not.

Insert Figure 4 about here

If one is willing to make two moderately reasonable assumptions it is possible to use the Bayesian Hierarchical Inference model to make predictions for the data in Fig. 4. The first assumption is that the participants all had .50-.50 prior probabilities about the success of each applicant before

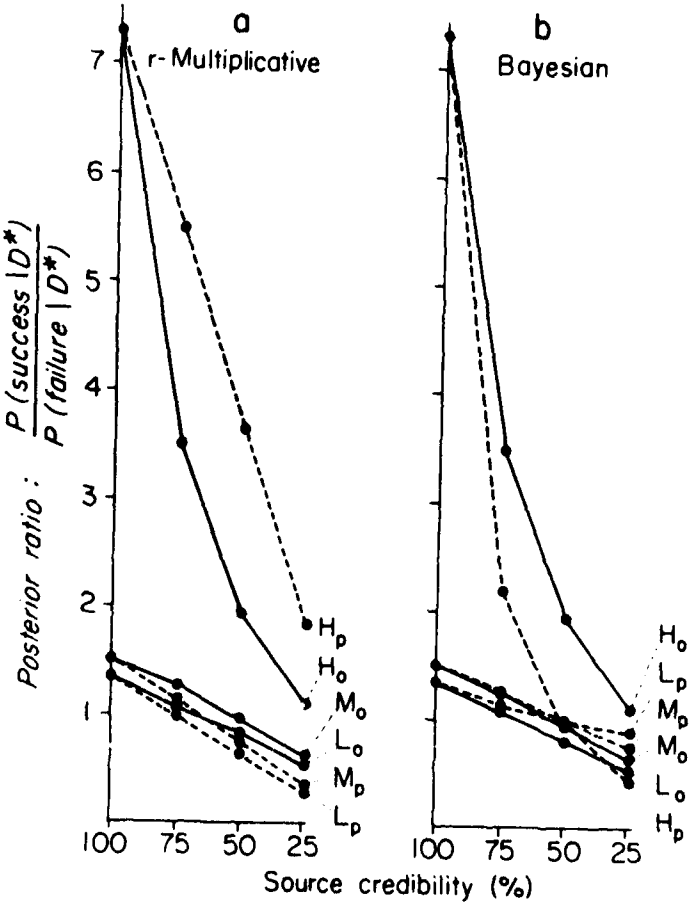

a
r-Multiplicative
b
Bayesian
Posterior ratio : P (success | D*) / P (failure | D*)
7
6
5
4
3
2
1
H_p
H_o
M_o
L_o
M_p
L_p
H_o
L_p
M_p
M_o
L_o
H_p
100 75 50 25
100 75 50 25
Source credibility (%)

Figure 4

they read the source's evaluation. The second is that in this experiment the probability of the report being correct, $P(D^*|D)$, is the same for a hit as for a correct rejection (i.e., the report on the applicant is properly positive or properly negative) and the probability of the report being wrong, $P(D^*|\bar{D})$, is the same for a false alarm as for a miss (i.e., the report is incorrectly positive or incorrectly negative). Using the same line of thought as was used to examine the multiplicative models, the first assumption means that the ratio of the participants' reported posterior probabilities for when the source was 100% credible, $\frac{P(H|D)}{P(\bar{H}|D)}$, are equal to their likelihood ratios, $\frac{P(D|H)}{P(D|\bar{H})}$; this is because the prior ratio was 1.00 and there was but one, undiscounted, revision. Of course, we do not know if these inferred likelihood ratios are accurate in a normative sense but that is not the point because, using the equation for Λ given in the introduction, the second assumption allows us to use them together with the 75%, 50%, and 25% source credibilities to compute what the posteriors ought to have been for each of these conditions if the participants were discounting the 100% likelihoods in a Bayesian manner. These predictions are compared with the observed probabilities (in ratio form) in panel b of Fig. 4.

At first glance it looks as though the Bayesian predictions may fit fairly well but this is not the case for two reasons. First, for 50% source credibility, the Bayesian model regards all information, whatever its relevance, as useless and the priors, likelihoods, and posteriors are all .50/.50 = 1.00; the participants did not do this even though it is intuitively reasonable, once one knows about it. The observed points, especially for high relevance are particularly divergent from the Bayesian prediction. Second, for the

Bayesian model the rate at which discounting occurs as source credibility decreases is different for different degrees of relevance (because of the differences in the initial sizes of their 100% credibility likelihoods). The predicted curve for the high relevance information drops quickly, that for medium relevance drops considerably less quickly, and that for low relevance information is almost flat. Indeed, the predicted ordering of the points for 25% source credibility is the reverse of that for 75%. More specifically, for 75% credibility, the Bayesian model predicts an order of high, medium and low relevance. For 25% credibility the order is low, medium and then high relevance. The curve for the observed data for the high relevance information drops quickly, but it does not drop far enough, and never reaches 1.00. And, while the other two curves for observed data are in the proper general area of the graph, they are roughly parallel rather than crossing as predicted and, as a result, the predicted difference in the ordering of the points for 25% and 75% credibility was not obtained.

Having rejected the subtractive, r-multiplicative, and Bayesian models we turn to an alternative. This is the Gettys et al. model with the r restriction removed. That is, r no longer just represents source credibility. It is adapted to represent two discounting factors, relevance and source credibility, with the values of the factors inferred from part of the data and used to predict the other part. The model is $P_a = P_p D_r D_s$, where P_a is the assessed probability of success for a hypothetical applicant, P_p is the posterior probability that would have been given for that applicant had the referral report contained completely relevant information from a source with 100% credibility, D_r is the discounting factor for the actual relevance of the information, and D_s is the discounting factor for the actual source credibility.

To assess the adequacy of the model, consider the case in which D_r and D_s are both 1.00. When the information is completely relevant and the source is 100% credible, $P_a = P_p$. Holding D_r at 1.00 the amount that P_a drops as D_s drops to 75%, or to 50%, or to 25% reveals the empirical values of D_s for each of these levels of reduced source credibility. Similarly, if D_s is held at 1.00 the amount that P_a drops as D_r drops to medium or to low reveals the empirical values of D_r for these levels of reduced relevance of the information. Then, having these values it is possible to use them in the above equation to predict P_a for all combinations of reduced (75%, 50%, 25%) source credibility and reduced (medium, low) information relevance--six combinations for each of the three jobs. The scatterplot of the predicted vs. observed P_a's is given in Fig. 5 and the inferred values of D_r and D_s for each job are given in Table 1. The model consistently under-predicts by an average of .02, but it clearly is in the ballpark.

Insert Figure 5 and Table 1 about here

Acceptability Condition. Unlike probabilities, ratings of acceptability have no "natural" scale and therefore care must be taken to avoid muddling the group data by combining ratings from participants who tend to limit their ratings to one or another parts of the scale. To avoid this, each participant's 36 ratings were Z-transformed using the mean and standard deviation of his or her own ratings.

Insert Figures 6 and 7 about here

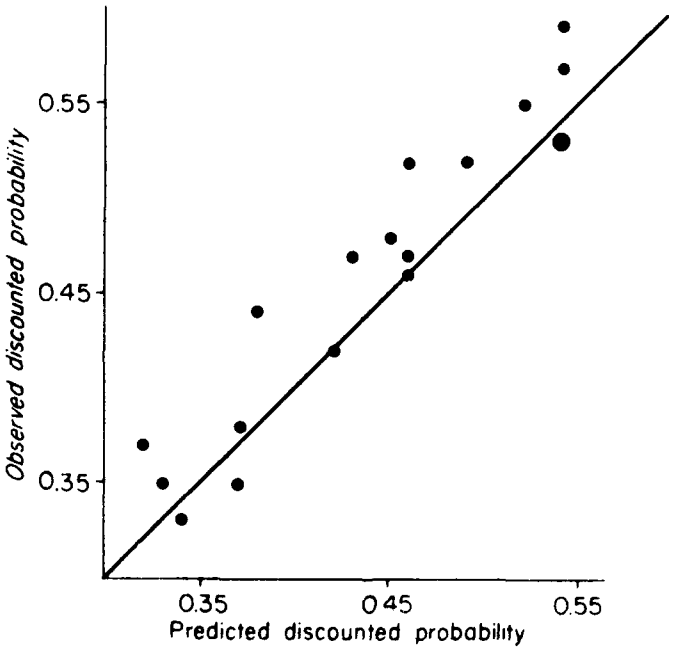

0.55
0.45
0.35
Observed discounted probability
0.35
0.45
0.55
Predicted discounted probability

Figure 5

Table 1

Inferred values of D_r and D_s for each job and across jobs

	D_r		D_s		
	M	L	75%	50%	25%
Driving Teacher	.72	.66	.86	.74	.60
Journalist	.67	.65	.90	.75	.56
Social Worker	.68	.68	.91	.77	.62
Mean	.69	.66	.89	.75	.59

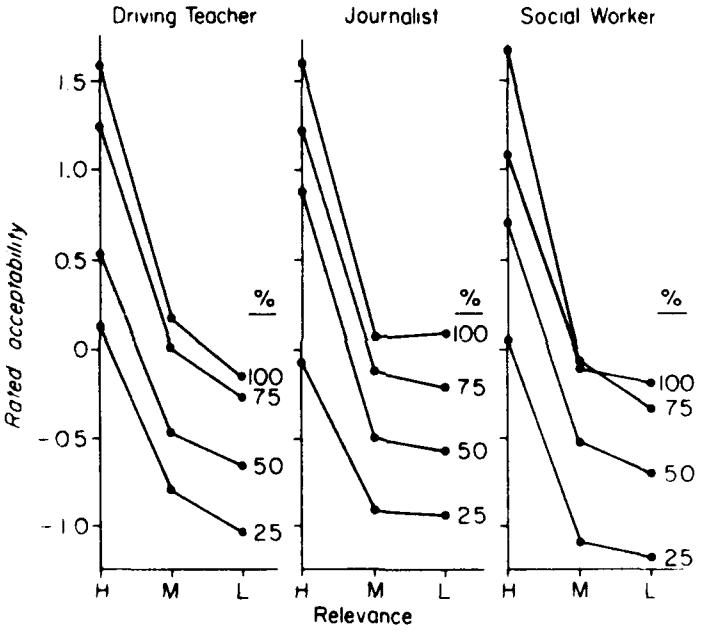

Driving Teacher
Journalist
Social Worker
Rated acceptability
1.5
1.0
0.5
0
-0.5
-1.0
%
100
75
50
25
H
M
L
Relevance

Figure 6

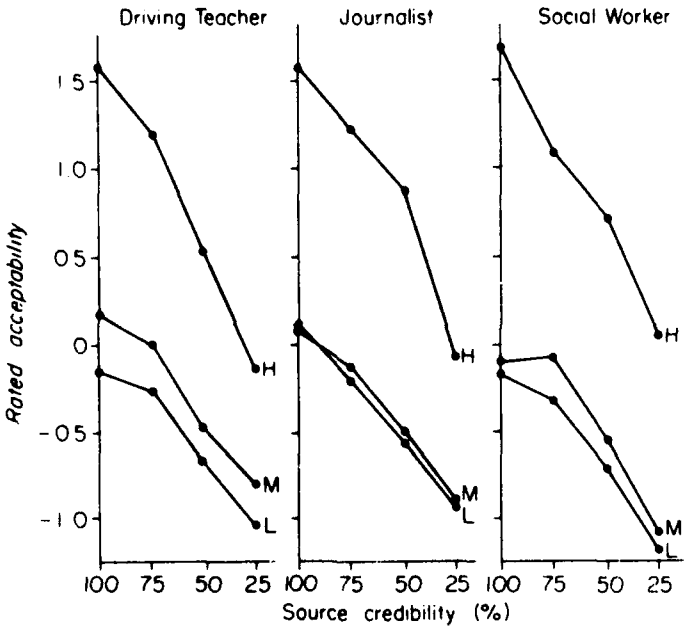

Driving Teacher
Journalist
Social Worker
Rated acceptability
1·5
1·0
0·5
0
-0·5
-1·0
H
M
L
100 75 50 25
Source credibility (%)

Figure 7

In Fig. 6 the mean Z-transformed acceptability ratings for each job are plotted with relevance as the independent variable and replotted in Fig. 7 with source credibility as the independent variable. There is a striking similarity between these graphs and Figs. 2 and 3. Looking at Fig. 7, the different elevations of the curves for the different levels of relevance indicate differential influence of these levels on rated acceptability. The decrease in the curves from left-to-right indicates discounting as a function of reduced source credibility. And, the curves appear non-parallel, indicating that discounting may be multiplicative rather than subtractive.

Analysis of variance confirms these observations. Jobs yielded no significant main effect, $F(2,70) = 1.54$, but relevance, $F(2,70) = 394.51$, $p < .001$, and source credibility, $F(3,105) = 153.44$, $p < .001$, were both significant as was the relevance x source credibility interaction, $F(6,210) = 10.42$, $p < .001$.

Again the log analysis was performed to make sure that the interaction could be accounted for solely by multiplicative discounting. Relevance was significant, $F(2,70) = 221.23$, $p < .001$, as was source credibility, $F(3,105) = 110.00$, $p < .001$, and their interaction was no longer significant, $F(6,210) = .44$.

The similarity between the results for these two different dependent variables, assessed probability and rated acceptability, is further supported by the correlation between their group means for each of the 36 hypothetical applicants, $r = .97$ ($p < .001$), using the Z-transformed acceptability data and $r = .98$ ($p < .001$), $a = .13$ ($p < .001$), $b = .86$ ($p < .001$) for the raw (non-transformed) data. The lack of complete identity between the two variables in the raw data analysis may be because of the "muddling" of scale-use biases, described above, resulting in less extreme raw data means.

Experiment 2

The first experiment answered the question about subtractive vs. multiplicative discounting, showed that in this kind of task the discounting factor was not equal to r, and showed a high degree of similarity between assessed probability and the evaluative variable, rated acceptability. However, the experiment was very simplistic in that it had only one piece of information and in that the information for each hypothetical job applicant was not varied, it was always favorable. Moreover, it is possible that by assigning an explicit numerical value to source credibility the experiment induced more deliberate and orderly discounting than might otherwise occur. Therefore, in this second experiment participants were asked to revise their opinions on the basis of three different pieces of information for each case. The implications of the information were varied and the source's credibility was qualitatively rather than quantitatively defined.

Method

The experimental materials consisted of "Social Impact Statements" for each of 216 hypothetical sites that were under consideration as possible locations for nuclear power plants. Each statement was on a separate sheet and contained three pieces of information, and a description of the information source. Each item of information referred to the potential negative impact of the proposed plant on "the social quality of life" in areas adjacent to the site. The statements varied in terms of three kinds of social impacts (Growth and Development, Local Institutions, Local Demography), three levels of probability that the negative impacts would indeed occur should the plant

be built (.75, .50, .25), and two levels of source credibility (high, low). Figure 8 contains an example of an impact statement for a site.

Insert Figure 8 about here

Participants were instructed to imagine themselves to be the member of the board of directors of an electric utility company. Their task was to weigh the social implications of possible siting decisions and to pass along to the board their well-considered judgments about how favorable each site was for construction of a nuclear power plant. Further, each site they were to review previously had been reviewed by the appropriate technicians and had been found to satisfy all technical, legal and ecological requirements. Therefore, the site was to be seen initially as wholly favorable and then favorableness was to be revised downward as a function of the three negative pieces of information in the social impact report. Participants were instructed to pay particular attention to the kind of information they were receiving, the message itself, and to the source of the information and how the information was obtained by the source.

Favorableness judgments were obtained on 100-point scales located at the bottom of each statement. The participant placed the pencil point on 100, read the first piece of information on the statement, revised downward, and marked the scale. Then he or she read the second piece of information and further revised and marked the scale, and then did the same thing for the third message.

The two levels of source credibility were determined before the experiment by presenting a large number of possible sources to a separate sample of

SITE NO .

THE FOLLOWING PREDICTIONS HAVE BEEN MADE CONCERNING MAJOR LONGTERM CONSTRUCTION AND OPERATION OF A NUCLEAR POWER PLANT ON THIS PARTICULAR SITE

IMPACT NO 1 CONSEQUENCES OF GROWTH AND DEVELOPMENT

ESTIMATED IMPACT — THE PROBABILITY IS 50 PERCENT OF A MAJOR LONGTERM IMPACT ON PRESENT REGIONAL PLANS FOR ECONOMIC GROWTH AND FOR LAND USE

SOURCE OF ESTIMATE — THE LONG-STANDING PREDICTION OF THE NATIONAL INSTITUTE OF SOCIAL AND ECONOMIC GEOGRAPHERS, SUBSTANTIATED BY MAPS, CHARTS, AND GRAPHS FROM RECENT LARGE-SCALE RESEARCH

IMPACT NO 2 INSTITUTIONAL CONSEQUENCES

ESTIMATED IMPACT — THE PROBABILITY IS 50 PERCENT OF A LONGTERM IMPACT ON THE QUALITY AND AVAILABILITY OF AREAWIDE SOCIAL AND MUNICIPAL SERVICES AND IN THE TYPES OF AGENCIES NEEDED TO ADMINISTER SUCH SERVICES

SOURCE OF ESTIMATE — A THREE-VOLUME REPORT, JOINTLY COMPILED BY THE AMERICAN INSTITUTE OF PLANNERS AND THE AMERICAN SOCIETY OF PLANNING OFFICIALS AS A SUMMARY OF THEIR EXHAUSTIVE STUDY OF MUNICIPAL INSTITUTIONS IN THE LOCALITY

IMPACT NO 3 DEMOGRAPHIC CONSEQUENCES

ESTIMATED IMPACT — THE PROBABILITY IS 25 PERCENT OF A MAJOR LONGTERM IMPACT ON THE DEMOGRAPHIC COMPOSITION OF THE LOCAL POPULATION (THAT IS, ON VITAL AND SOCIAL STATISTICS SUCH AS NUMBERS OF BIRTHS, DEATHS, MARRIAGES, ET CETERA)

SOURCE OF ESTIMATE — SYSTEMATIC STUDIES MADE BY DEMOGRAPHIC EXPERTS WHICH WERE BASED ON ACCURATE POPULATION DATA AND EXTENSIVE SURVEYS

MY ESTIMATE OF THE OVERALL FAVORABILITY OF THIS SITE

0 10 20 30 40 50 60 70 80 90 100

LOW FAVORABILITY — HIGH FAVORABILITY

Figure 8

persons. Sources were selected that were consistently rated either high or low.

Because the task took quite a while to complete, participants worked on it for an hour or so on successive days until they completed all statements. There were 29 participants who were college students and who were paid $15 upon completion of the task.

There are six possible ways to order three items of information. The three kinds of social impacts, the three probabilities associated with each, and the two levels of source credibility were presented in all possible orders which resulted in 216 separate impact statements.

Predictions. On the basis of the results of the first experiment it was predicted that the amount of revision would be larger for high credibility sources than for low credibility sources. The amount of revision should decrease as a function of decreases in the probabilities of the negative impacts actually occurring. Additionally, because a large number of subjective probability revision studies have found the first revision to be larger than subsequent revisions even when the information does not warrant decreasing revision and in accord with our hypothesis that the same mechanisms may underlie various kinds of opinion revision, it was predicted that the first of the three revisions of each social impact statement would be larger than the subsequent two, even when the source credibility and message content were constant for all three revisions.

Results

The difference, D, between successive marks on the favorableness scales was the dependent variable. There are three of these differences for each of

the 216 impact statements, making a total of 648 for each participant. To eliminate the effects of response biases (some persons tended to revise a lot and others only a little) on group data, each participant's D's were Z-transformed using the mean and standard deviation of his or her own set of D's.

Because 648 data points for each participant requires an unwieldly analysis, the data were reduced by averaging across presentation order for the three kinds of impacts, three levels of probability, and the two levels of source credibility, 3 x 3 x 2 = 18. The resulting 18 mean data points for each participant were then averaged across the group for the data analysis.

Figure 9 shows the mean degree of favorableness revision as a function of source credibility for each level of probability of the impact's occurrence for each of the three kinds of social impacts. Clearly the degree of downward

---

Insert Figures 9 and 10 about here

---

revision, D, of the favorableness of a site is a positive function both of the probability that a negative social impact will actually occur if a plant is built on it, and the credibility of the source that reports the probability, and in addition, it is different for different kinds of social impacts. Moreover, the lines in the graph appear nonparallel, implying that information discounting was multiplicative.

An analysis of variance yielded significant main effects for kinds of social impacts, $F(2,56) = 8.01$, $p < .001$, levels of probability $F(2,56) = 357.12$, $p < .001$, and source credibility, $F(3,84) = 62.36$, $p < .001$. There was a significant interaction between kinds of impacts and levels of probability, $F(4,112) = 10.85$, $p < .001$, which in Fig. 9 is reflected in the varying

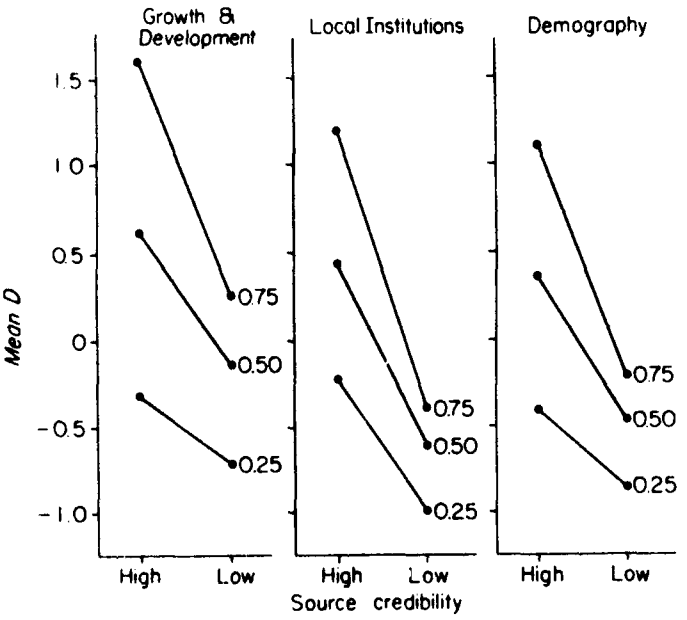

Growth & Development
Local Institutions
Demography
Mean D
1.5
1.0
0.5
0
-0.5
-1.0
0.75
0.50
0.25
High
Low
Source credibility

Figure 9

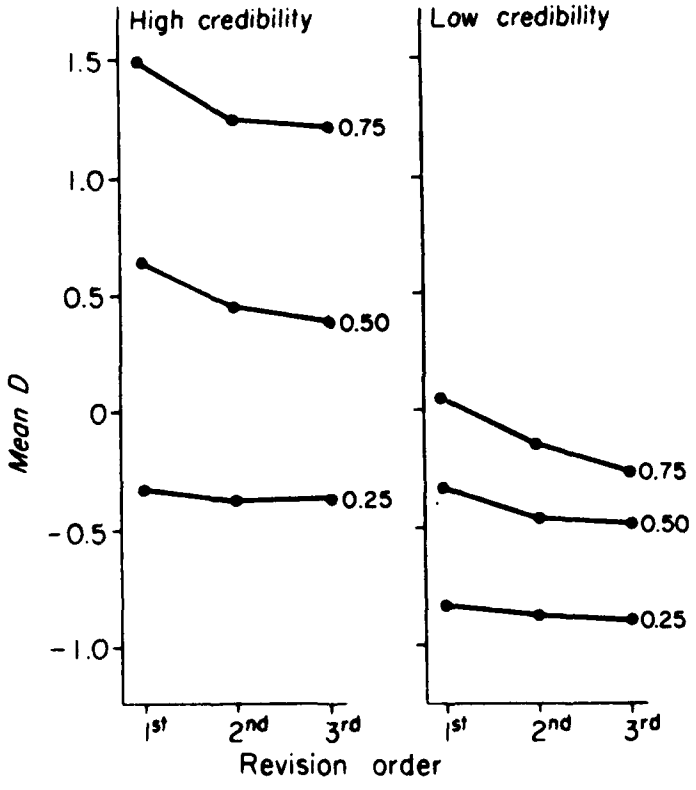

High credibility
Low credibility
1.5
1.0
0.5
0
-0.5
-1.0
Mean D
0.75
0.50
0.25
0.75
0.50
0.25
1st
2nd
3rd
1st
2nd
3rd
Revision order

Figure 10

heights of the lines for the different kinds of social impacts. This interaction is of *no* substantive importance. However, the critical result, the interaction between levels of probability and source credibility was significant, $F(6,168) = 61.00$, $p < .001$, indicating that, here too, discounting was multiplicative.

As in the previous analyses, the data were reanalyzed under a log transformation to see if the probability x source credibility interaction was solely attributable to multiplicative discounting. The three main effects remained significant; for kind of impact $F(2,56) = 8.67$, $p < .001$, for levels of probability $F(2,56) = 160.36$, $p < .001$, and for source credibility $F(3,84) = 53.28$, $p < .001$. The kind of social impact x probability interaction was non-significant, $F(4,112) = .71$. The important interaction between probability and source reliability remained significant, $F(6,168) = 2.90$, $p < .05$, but the proportion of variance accounted for by the interaction in the two analyses goes from .05 in the first analysis to .004 in the log analysis, a substantial reduction. For all intents and purposes discounting was accounted for by a multiplicative process.

To examine the magnitudes of revisions as a function of their being the first, second, or third revision, each participant's D's were averaged across kinds of social impacts for each of the three revisions. The three levels of revision order, the two levels of source credibility, and the three levels of probability yield 3 x 2 x 3 = 18 mean data points for each of the 29 participants, which were then averaged across the group.

Fig. 10 shows the mean degree of favorableness revision as a function of revision order for each level of probability for both high and low credibility

sources. The first revision appears to be greater than the second which is, with the exception of .25 for high credibility sources, greater than the third. The differences, in some cases, are small but they are significant. An analysis of variance yielded significant main effects for revision order, $F(2,56) = 10.2$, $p < .001$, source credibility, $F(1,28) = 70$, $p < .001$, and probability $F(2,56) = 376$, $p < .001$. There was a significant interaction between source credibility and probability, $F(2,56) = 116.6$, $p < .001$, which was also seen in the previous analysis and that is represented by the differing heights of the curves in the two graphs in Fig. 10. There also was a significant interaction between revision order and probability, $F(4,112) = 11.6$, $p < .001$ that can be seen in the figure as a decreasing effect of revision order on D as probability decreases.

Discussion

The results of both experiments clearly show that, at least for these kinds of laboratory tasks, discounting of information in opinion revision is heavily influenced by relevance of the information and by the credibility of its source. This discounting process seems to be multiplicative, not subtractive. Moreover, results for the probability variable in the first experiment show that the discounting factor was not source credibility, r, which disconfirms the Snapper and Fryback (1971) model, although the model easily could be altered to omit this restriction. The Gettys, Kelly, and Peterson (1973) model is not disconfirmed by these results because it does not demand that r be equal to source credibility and indeed it was possible to fit part of the data with this model by inferring the necessary discount factors from another part of the data. Unless future research can find a systematic link between

the size of the discounting factor that people use and reasonable situational variables, particularly source credibility, it means that the factor will always have to be empirically inferred from the data for each application of the model--a requirement that makes the model less valuable than could be desired. (This criticism also would apply to a liberalized version of the Snapper and Fryback model.) Note, however, that the present results do not speak to the other difference between the two models--whether it is the likelihood ratio or the posterior probability ratio that is discounted, although in using the Gettys et al. model we assumed that it was the latter. This is a potentially important practical and theoretical point that deserves further attention.

The second experiment demonstrates that source credibility need not be quantitatively stated for its effects to be seen in opinion revision. Also, when a series of revisions is required for a single hypothesis the amount of revision decreases, albeit only slightly, from one revision to the next even though objectively, the implications of the information remain constant. It is not clear whether this decrease is 'conservatism' of Bayesian fame, whether it is merely some sort of anchoring effect as discussed by Tversky and Kahneman (1974), or whether they all three are the same thing. But having obtained it in a revision experiment in which subjective probability was not the dependent variable means that any explanation cannot be limited to probability revisions alone; the explanation must be general enough to encompass other kinds of opinion revisions as well. Indeed, the similarity in results between acceptability and probability response variables in Experiment 1 and between the conservatism result in Experiment 2 and the large literature on that topic, suggests

that subjective probability revision and other kinds of opinion revisions have more in common than just similar mechanisms--they may be exactly the same thing. If there is identity and if revisions do not follow the laws of probability, for which there is ample evidence (Slovic & Lichtenstein, 1971, Tversky & Kahneman, 1974), then the special place of normative probability theory in decision psychology becomes precarious. If there is identity and if revisions of subjective uncertainty play a role in unaided decision making, even if they do not conform to the dictates of probability theory, it may prove profitable to broaden the investigation of this uncertainty, and related opinions, by linking decision research with attitude research. This linkage might put a little life back into both areas of endeavor and lead to the development of a comprehensive theory of opinion and opinion revison.

References

Anderson, N. H. Integration theory and attitude change. Psychological Review, 1971, 78, 171-206.

Choo, T. Communicator credibility and communication discrepancy as determinants of opinion chnage. Journal of Social Psychology, 1964, 64, 65-76.

Gettys, C. F., Kelly, C. & Peterson, C. R. The best guess hypothesis in multistage inference. Organizational Behavior and Human Performance, 1973, 10, 364-373.

Gettys, C. F. & Willke, T. A. The application of Bayes' theorem when the true data state is unknown. Organizational Behavior and Human Performance, 1969, 4, 125-141.

Hill, A. H. Credibility, discrepancy and latitude of communication as dimensions of dissonance influencing attitude change. Australian Journal of Psychology, 1963, 15, 124-132.

Hovland, C. I., Janis, I. L. & Kelley, H. H. A summary of experimental studies of opinion change. In Jahoda, M. & Warren, N. (Eds.) Attitudes, Baltimore: Penguin, 1966, 139-151.

Peterson, C. R. Introduction to special issue on hierarchical inference. Organizational Behavior and Human Performance, 1973, 10, 315-317.

Schum, D. A. & DuCharme, W. M. Comments on the relationship between the impact and the reliability of evidence. Organizational Behavior and Human Performance, 1971, 6, 111-131.

Slovic, P. & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Snapper, K. & Fryback, D. G. Inferences based on unreliable reports. Journal of Experimental Psychology, 1971, 87, 401-404.

Tversky, A. & Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.

Footnotes

[1]This research was partially supported by Battelle Institute Northwest, Project Number BSA 936 and by Office of Naval Research Contract N00014-76-C-0193.

The authors wish to thank Mr. Stephen Lawrence, Mr. Jeffrey Prothero, and Drs. Barbara H. Beach, David A. Schum and John W. Prothero for their help.

[2]Requests for reprints should be sent to Dr. L. R. Beach, Department of Psychology, NI-25, University of Washington, Seattle, WA 98195.

Figure Captions

Figure 1. An example of a "Referral Report" for Experiment 1.

Figure 2. Mean assessed probability of success as a function of information relevance for each level of source credibility for each job.

Figure 3. Mean assessed probability of success as a function of source credibility for each level of information relevance for each job.

Figure 4. Predicted (p) and observed (o) posterior probability ratios as a function of source credibility for each level of information relevance (H, M, L) for the r-multiplicative models and the Bayesian model.

Figure 5. Predicted and observed discounted posterior probabilities for the Gettys et al. model with the r restriction removed.

Figure 6. Rated acceptability as a function of information relevance for each level of source credibility for each job.

Figure 7. Rated acceptability as a function of source credibility for each level of information relevance for each job.

Figure 8. Example of a "Social Impact Statement" for Experiment 2.

Figure 9. Mean D as a function of source credibility for each level of probability of a negative social impact for each kind of social impact.

Figure 10. Mean D as a function of whether the revision was first, second or third (revision order) for each level of probability of a negative social impact for high and for low source credibility.